

In preparation for *Mind and Language*

THE CAUSAL INEFFICACY OF CONTENT

GABRIEL M. A. SEGAL

ABSTRACT: The paper begins with the assumption that psychological event tokens are identical to or constituted from physical events. It then articulates a familiar apparent problem concerning the causal role of psychological properties. If they do not reduce to physical properties, then either they must be epiphenomenal or any effects they cause must also be caused by physical properties hence be overdetermined. It then argues that both epiphenomenalism and overdeterminationism are *prima facie* perfectly reasonable and relatively unproblematic views. The paper proceeds to argue against Kim's (Kim 2000, 2005) attempt to articulate a plausible version of reductionism. It is then argued that psychological properties, along with paradigmatically causally efficacious macro-properties, such as toughness, are causally inefficacious in respect of their possessor's typical effects, because they are insufficiently distinct from those effects. It is finally suggested that the distinction between epiphenomenalism and overdeterminationism maybe more terminological than real.

1. Introduction

I begin with some territory that will be familiar to most readers. Here is how the problem of mental causation has typically been set up since shortly after the onset of

non-reductive physicalism.¹ It is now widely assumed that the realm of the physical is causally closed. This means that the probability of any event's occurring is fully determined by physical causes. This apparently leaves us with a limited number of options concerning psychological causation, none of which appear hugely attractive. Either: (a) the psychological is epiphenomenal and can have no causal impact on the physical, or (b) the psychological is identical with the physical, or (c) actions are all overdetermined, each one having at least two distinct sufficient causes. Option (b) subdivides into two further options. Either (b1) the psychological reduces to the physical and every psychological property is identical with some physical property, or

Many thanks for comments and discussion to Helen Beebee, Tim Crane, Christopher Hughes, Keith Hossack, Jaegwon Kim, Cynthia MacDonald, Sydney Shoemaker, David Papineau, Elliott Sober, Robert Templing, Ann Whittle and David Yates.

Address for correspondence: Department of Philosophy, King's College, Strand, London WC2R 2LS

Email: Gabriel.Segal@kcl.ac.uk

¹ Cartesian dualism, which holds that minds and psychological events are non-physical, is generally held to be inconsistent with the existence of causal interaction between the psychological and the physical. Indeed, Descartes's contemporaries saw causation as a problem for him: 'I could more readily allow that the soul has matter and extension than that an immaterial being has the capacity of moving a body and being affected by it' said Princess Elizabeth of Bohemia (letter of June 1643, translated in Anscombe and Geach eds., 1971). Descartes's reply to Elizabeth is widely agreed to be unconvincing. However, I think that given what was known at the time, Descartes would have been able to offer a reasonable defence of his position. He might have adopted any of various accounts of causation (e.g. appealing to constant conjunction or to counterfactuals) and argued that the chosen account ratifies the existence of psychophysical causation without entailing that the mind is extended. The problem for Cartesian dualism only became a matter of clear and serious concern in the mid 1960s, by which time neurology had advanced to a point where it became reasonably clear that there is a complete physical causal explanation for every human action (see Papineau, 2000 for discussion). And it is this that creates the current problematic.

(b2) token psychological events are identical with or constituted from token physical events but psychological properties are not identical with physical properties. (b1) is widely held to be inconsistent with the multiple realization of the psychological by the physical. And (b2) appears to bring us back to the original problematic, with the properties as the locus of tension. If one event causes another, it does so in virtue of some of its properties and not others. If I throw a stone at a window and the window breaks, it is because the stone was hard and heavy that it broke the window and not, say, because it was grey and millions of years old. The properties in virtue of which an event has a particular effect are typically called the ‘causally efficacious properties of the cause with respect to the effect.’ Suppose, then, that a token neural event causes an action. We can ask ‘Does it do so in virtue of its physical properties or its psychological properties?’ and we are back to choosing between options (a) and (c) or returning to (b1). And none of those options appeal.

That, as I say, is how the problematic is typically set up within the Davidson-Fodor package of token identity (or constitution), type dualism (or type pluralism) and supervenience (e.g. Davidson, 1970, Fodor, 1974). I will call this, ‘The Standard Package’. I think there is quite a lot that is seriously questionable about that way of looking at things. I want to accept The Standard Package and take a look at the problematic. My plan is as follows. First, I will say a few words about why I think that neither (a), epiphenomenalism, nor (c), overdetermination, should be rejected as quickly as they often are. Indeed, these options are sometimes dismissed so quickly that it is difficult to find articulated arguments against them. Kim (2000, 2005) does articulate some arguments, and I will use them as my focus. Kim’s rejection of (a) and (c) lead him to favour a form of (b), reductionism. I will go on to criticise Kim’s reductive approach. I will then turn to a proposal due to Segal and Sober (1991) who

defend a version of (c), overdetermination, and argue that the proposal does not work. I will then go on to argue for a version of (a), epiphenomenalism. Finally I will suggest that the distinction between overdeterminationism and epiphenomenalism may be more rhetorical than real.

2. The Standard Package and Overdetermination

Let me begin with The Standard Package - in particular Fodor's version from Fodor (1974) - and why it might seem fitting to add overdetermination to it. Let us suppose that we have a (non-strict) psychological law to the effect that every event with psychological property F causes an event with psychological property G; 'Ceteris paribus, every time someone has a strong desire to jump up and down, they jump up and down,' for example. F has a wide variety of possible physical supervenience bases, $m(F)_1, \dots, m(F)_n$. Every time the law is instantiated, the occurrence of the F event is realised by the occurrence of some $m(F)$ event. The $m(F)$ event causes, by physically explicable, unmysterious mechanisms, the occurrence of some $m(G)$ event, where $m(G)$ is a supervenience base for G.

Let us for the time being just accept that metaphysical picture. Then it is tempting to say that just as paradigmatic psychological properties (e.g. being a desire to jump up and down) supervene on physical properties, so also psychological² laws supervene on physical laws and psychological causal relations supervene on physical causal relations. And, if we are tempted to say that, then we might well be tempted by overdetermination. For we do not conclude from the fact that psychological properties

² I will use 'psychological' as an abbreviation of 'psychological and psychophysical' when talking of laws and causation.

supervene on physical ones that psychological properties don't exist. We do not conclude from the fact that psychological laws supervene on physical laws, that there are no psychological laws. So, to be consistent, we should adopt a similar attitude to psychological causation: we should not conclude from the fact that psychological causation supervenes on physical causation that it does not so much as occur. We have independent reason to believe that the m(F)s are causally efficacious in respect of causing G events. But that should not prevent us from allowing that F itself is also causally efficacious in respect of the very same effects. We should not think of F and the m(F)s as competing for efficacy. Rather, we should think of F's efficacy as derived from its m(F)s.³

So, on this approach, both m(F) and F are causally efficacious in the matter of causing Gs. That means that they are both sufficient for G events, under the circumstances. Some regard this idea as highly problematic. Kim, for example, regards 'the overdetermination approach' as 'a non-starter' (Kim 2000). But his arguments against it hinge crucially on a principle of *Exclusion* (2005: 42)

Exclusion No single event can have more than one sufficient cause occurring at a given time – unless it is a genuine case of causal overdetermination.

In Kim's terminology, a 'genuine' case of causal overdetermination is one in which 'each overdetermining cause plays a distinct and distinctive causal role' (2005 42).

Kim doesn't argue for *Exclusion*. He does offer remarks such as the following (2005 48):

³ This is Segal and Sober's (1991) position as well as, arguably, that of Kim (1984). See also Sider (2003) for friendly remarks about overdetermination.

To be a cause of [m(G)], [F] must somehow ride piggyback on physical causal chains ... And we may ask in virtue of what relation it bears to physical property [m(F)] does [F] earn its entitlement to a free ride on the causal chain from [m(F)] to [m(G)] and to claim this causal chain to be its own?

I find it hard to interpret the question. The question of whether F is causally efficacious in bringing about m(G) is not a question of rights. Consider the analogy with existence. 'To exist, F must somehow ride piggyback on a physical property m(F) and we may ask in virtue of what does F earn its entitlement to free ride on m(F) and to claim its existence to be its own?'

The question carries no force. The thing to do is not to prejudge the issue but to look and see whether supervenient properties exist and whether they are efficacious in bringing about effects. And on first inspection, it looks as though they do. That is the way the world appears to us empirically. For example, it appears to us as though stones break windows, in part, because they are tough.⁴ Toughness supervenes on but does not reduce to microstructure. And properties like toughness (or at least analogues thereof, such as fracture toughness and impact strength) are real measurable properties understood in depth and detail by material science.

Kim thinks that *Exclusion* is obviously true: 'I believe it is virtually an analytic truth with not much content' (2001: 51). I take it that if *Exclusion* is virtually analytic then it is not actually analytic. But then, in the face of appearances, the onus

⁴ The appearance is given some support by the counterfactual: if the stone hadn't been tough, it wouldn't have broken the window. For more on counterfactuals and mental causation see Bennett (2003).

is on its proponents to provide some argument for it. The empirical trumps the almost analytic.

Of course, if it were merely a coincidence that every time an F event caused a G event, some m(F) event or other just happened to crop up, for no apparent reason, and cause that same G event, then that would indeed be incredible. That would be analogous to a world in which every time somebody was killed by a bullet they were also, at the same time, killed by some other unrelated means. But that is not how it goes with the standard package. The occurrence of F events and m(F) events are, of course, not unrelated. Rather, the co-occurrence of Fs and m(F)s is guaranteed by the supervenience component of the package.

On this view, overdetermination is not rare, but ubiquitous. It holds in all cases that fit the metaphysics of The Standard Package. If I throw a stone at a window and the window breaks, then we can explain what happened at more than one level. For example, we can explain it in relatively macroscopic terms, appealing to things like the stone's toughness and mass. And we can explain it at the microscopic level by appealing to things like the arrangement of the stone's component atoms and the strength of bonds among them. Both explanations appear to cite causally efficacious properties of the stone: e.g. its being tough and its being so-composed of atoms. The explanations don't compete. So why suppose that the properties compete for the title of 'causally efficacious'? Don't prejudge the question of whether overdetermination is ubiquitous, but look and see whether it is.

It is fruitful (following Yablo 1992) to think about the issue in relation to determinates and determinables. Suppose we place a crimson cube in front of a mirror. This causes an image to appear to in the mirror. The image has many properties. It is crimson and, let's say, square. But it is also red and rectangular.

Now what properties of the object caused there to be a red, rectangular image?

One obvious answer is: redness and rectangularity. But another perfectly good answer is: crimsonness and squareness. These properties of the object cause there to be a crimson square image. And, since it follows logically that the image is also red and rectangular, they cause there to be a red, rectangular image. There is nothing very mysterious about this. Both crimsonness and redness, for example, have the power to cause red images, and the reason they do their causing at the same time and in relation to the same effects is because being crimson is a way of being red. Moreover, each time a red thing causes a red image, the redness is realised by some determinate shade that causes a correspondingly shaded image. But there is nothing particularly mysterious about that either. There are just many ways of being red and every red thing has got to be red in some way or other.⁵

Overdeterminationism, then, is at least a starter. What, then, of epiphenomenalism?

3. Kim's Epiphobia and Reductionism

Epiphenomenalism about psychological properties has had only a small number of champions (e.g. Jackson 1982). And there is something of a tendency for people to dismiss it too quickly. Some think that the causal efficacy of psychological properties is given in our experience. Surely I know why I jumped up and down: it's because I wanted to.

⁵ Yablo argues that the kind of supervenience relation that holds between mental and physical is actually a variety of the determinable-determinate relationship.

It is true that some of the time, it does seem to us that we know why we do something. I might contemplate whether to jump up and down, be fully aware of my desire to do so and then, in the light of this desire, go ahead and jump. In such a case, it might well seem to me that my desire to jump up and down causes me to jump up and down. But, in my case at least, I don't feel that I have a particular sense of which properties of that desire caused the action. Nothing in my experience contradicts the idea that it was its physical properties, and those alone, that did the work.

Kim cites a case that might seem more compelling: pain. Suppose that I suddenly feel a burning pain, as I accidentally lean on the hob. I instantly pull my hand away. Surely, someone might say, it is because your hand hurt that you moved it. But even in that case, I don't think my experience presents me with a view about which properties are doing the causal work. I feel the pain and I move my hand, but I do not experience the painfulness of the pain causing the hand to move. It is Hume all the way.

In fact, both experience and science offer some support to epiphenomenalism about conscious states. For sometimes we feel the pain after we have performed the action that we might otherwise think was motivated by it. And science tells us that sometimes the neural events that initiate the causal processes of our actions occur before we know what we are trying to do.⁶

Many people also seem to be motivated by the idea that it would be unfortunate - even terrible - if epiphenomenalism were true, and that some of our treasured views about ourselves and our place in nature would be under threat. Kim

⁶ See e.g. Libet (1985), and Haggard and Eimer (1999). The data certainly don't prove epiphenomenalism. But they do suggest that the brain decides what the subject is going to do and then afterwards informs the subject of its decision.

gives voice to three concerns of that sort. If we keep clearly in mind that it is the causal potency of psychological properties that is at issue, I think that we can see that the concerns are groundless.

First (2000 p. 31): ‘the possibility of human agency evidently requires that our psychological states ... have causal effects in the physical world’. A little clarification is in order here. It does seem right that my agency requires that I do things because of what I want, what I believe and so on. And it seems right to take that as a causal ‘because’. So, for example, if I jump up and down just because I want to, it is my desire to jump up and down that causes me to jump up and down. But that is token-event causation. What is not at all evident is that our conception of human agency requires further that the psychological properties of my desire be causally efficacious in moving my body.

Second (ibid.): the possibility of human knowledge presupposes the reality of psychological causation. For example, Kim says, reasoning requires the causation of beliefs by beliefs and memory requires the causation of beliefs by perceptions. I don’t think that that’s at all plausible. The acquisition of knowledge is compatible with epiphenomenalism even about token mental events. This is not the place to get involved with detailed epistemology. But, on standard views, what is required for knowledge is that beliefs be caused in ‘the right way’, and whatever that ‘right way’ is, it doesn’t require the causes to be psychological. That is part of the point of automated computation. In a computer, physical processes ensure the production of representations that stand in appropriate rational relations to one another, given their interpretations. If we are ensembles of computers, as cognitive science supposes we are, then we could be excellent knowledge-acquisition devices whether or not our representations’ psychological properties have causal potency. Suppose, for example,

that one of a certain creature's belief-forming mechanisms is a physically implemented truth-preserving inference machine that only creates a new belief if it is a logical consequence of old ones. If other conditions are right, then this creature would be in a good position to acquire new pieces of knowledge via the operation of the mechanism, epiphenomenalism or no.

Third (still *ibid.*): 'the possibility of psychology as a theoretical science capable of generating law-based explanations of human behaviour depends on the reality of psychological causation: psychological phenomena must be capable of functioning as indispensable links in causal chains leading to physical behaviour'. Again, this might be right if we are talking of psychological events and processes. Many existing branches of psychology are committed to the existence of causal psychological laws. The laws relate psychological phenomena to their effects. But these sciences are not committed to telling a complete story about which properties of the causes are responsible for bringing the effects about. This is clear, for example, in classical cognitive science. In theorising about a given cognitive system, a cognitive scientist aims to tell us three stories about which representations will cause the production of which other representations: semantic, syntactic and physical (e.g. neural). *Prima facie*, given the way most of us tend to think about causation, it looks as though cognitive science does aim to tell us something about which properties of representations are causally responsible for their characteristic effects: the physical ones. (I am not convinced that even that much is right. But let that pass.) But what cognitive science emphatically does not tell us is whether the syntactic and semantic properties are causally efficacious in respect of the same effects. That is why the matter is still debated by philosophers. Existing psychology leaves open questions about the efficacy of psychological properties. And future psychology will be able to

accommodate the causal role of psychological properties, whatever that role turns out to be.

Kim's rejection of overdetermination and his epiphobia lead him to try and formulate a plausible version of reductionism. Like most of us, Kim accepts multiple realisation. The proposal he outlines (which is similar to David Lewis's (Lewis 1980) brand of functionalism), is as follows. (Kim does not fully endorse the proposal but rather expresses hope that it might turn out to be right). On the view Kim proposes, psychological descriptions are equivalent to second-order functional descriptions.⁷ So, for example, to satisfy some mental predicate 'M' is to have some first-order property, P_n , such that P_n has some particular characteristic causal role. P_n will be a physical property, in effect a realiser (an $m(M)$) for the mental property M. Kim assumes that M has multiple physical realizers in different 'species and structures and can have different realizers in different possible worlds' (2001 110). He continues (ibid.):

The reduction consists in identifying M with its realizer P_i relative to the species or structure under consideration (also relative to the reference world). Thus M is P_1 in species 1, P_2 in species 2, and so on. Given that each instance of M has exactly the causal powers of its realiser on that occasion ... all the causal explanatory work done by an instance of M that occurs in virtue of the instantiation of realizer P_1 is done by P_1 ...

⁷ Kim doubts that this is true of terms for qualitative mental states, but thinks it might be true of representational states.

This offers an illusion of saving the efficacy of mental properties. Here is the illusion: we have a mental property M , in a species or structure k , at a possible world W_j , and that property is identical with some P_k and so, unproblematically, has P_k 's causal powers. A couple of clarifications should bring out why it is an illusion. First, 'species or structures' are those items that have just one way of realising M at any given world. Maybe talk of species is here appropriate for some types of mental property, like, say, the sort of pain caused by burning. (Note, however, that Kim himself is not optimistic about the possibility of the story applying to phenomenal properties). And it might be appropriate for certain representational properties as well, such as representations in hard-wired bits of computational modules. For such things, maybe, M has a single realizer across a whole species. But talk of species is not appropriate for lots of interesting psychological states. Consider propositional attitudes, for example. It is very unlikely that a given propositional attitude, such as the belief that Barcelona is beautiful, will have the same physical realisations in different people, or, perhaps, in the same individual at different times. So, at least as far as propositional attitudes are concerned, 'structures' are likely to be individual cognizers, or cognizers-during-periods, at particular worlds.

The second point of clarification is that what get identified with the P_n s are not psychological properties of the kind whose efficacy was in question. The properties at issue are those ascribed by psychology, like the property of believing that Barcelona is beautiful. What get identified with the P_n s are, rather, properties that might be described along the lines of: M -in-structure- K -in- W_j . And of course M -in-structure- K -in- W_j , M -in-structure- K -in- W_k and M -in-structure- J -in- W_l etc. are all different properties. But such properties are not the ones that feature in psychological generalisations. If you and I and Genoveva all believe that Barcelona is beautiful, then

we share an interesting psychological property and the question arises of whether that property is efficacious. It does not help answer that question to say that the property of being my-actual-belief-that-Barcelona-is-beautiful is identical with some P_i and therefore is efficacious and the property of being your-actual-belief-that-Barcelona-is-beautiful is identical with some P_j and so it, too, is efficacious, etc..

The illusion is thus that there exists a tenable position that saves mental causation via a version of type-identity theory, while at the same time endorsing multiple realisation. Unfortunately, there isn't. The properties in whose efficacy we are interested are the multiply realised ones, and those cannot be identified with the physical.

In fact, psychological properties suffer a worse fate than mere impotence, in Kim's own opinion. Psychological descriptions are to be reconstructed as functional descriptions, so, at first pass, we would construe psychological properties as functional properties. But Kim doesn't really believe in functional properties, since he makes a case for 'eschewing the talk of functional *properties* in favor of functional *concepts* and *expressions*.' (Kim 2000, 110.) According to Kim, functional concepts and expressions are perfectly good and useful. But they don't pick out real properties: or at least they don't pick out 'robust' properties of the sort that feature in proper scientific generalisations.

I think that the fate of Kim's package reveals the dangers of believing in multiple realization while rejecting both overdetermination and epiphenomenalism. It very much looks as though multiple realization is just one of those things we will have to accept. So it seems sensible to see if we can make a reasonable case for one of the other options. I will now outline Segal and Sober's (1991) defence of an

overdetermination approach, and argue that it fails. I will then offer an alternative that I will present as a version of epiphenomenalism.

4. Segal and Sober

Segal and Sober offer a sufficient condition for the causal efficacy of a macro-property, which they label '(P5)', and argue that psychological properties meet that condition. I will state (P5), which is a touch complicated, and say a little about the motivation behind it.

(P5) If (i) it is a (possibly nonstrict) law that every F event causes a G event and (ii) in each case in which an F event causes a G event there exist micro-properties $m(F)$, $m(F)'$ and $m(G)$ such that the cause's being F mereologically supervenes on its being $m(F)$ and the effect's being G supervenes on its being $m(G)$ and possession of $m(F)$ includes possession of $m(F)'$ and the cause's being $m(F)'$ causes the effect's being $m(G)$, then F is efficacious in the production of Gs.

(P5) is an attempt to explain macro-causation in terms of micro-causation. The idea is that the combination of lawfulness and supervenience requirements should be strong enough to rule out all counterexamples. It is well known that laws aren't enough: for example, successful matings of blue-eyed individuals cause births of blue-eyed children. But the property of involving blue-eyed individuals is not causally efficacious in respect of the mating's production of blue-eyed children. Rather, it's a property relating to the parents' genes that causes both the parents and the children to have blue eyes. Segal and Sober's thought was that correlations of this sort could be

ruled out by requiring a sufficiently tight relation between the macro-property of the cause and the micro-properties that we assume to be doing causal work in bringing about the effect. Mereological supervenience was thought to do the trick. Roughly speaking, mereological supervenience is the converse of the ‘makes it the case that’ relation that we talk about when we explain why an object has certain macro-properties by citing properties of and relations among its components (at some level of description): e.g. the way the diamond’s crystals are bound together makes it the case that the diamond is hard. The property of being blue-eyed does not mereologically supervene on the relevant genetic properties of a successful mating between blue-eyed individuals (although it does supervene on them in some weak sense).

In fact, even by Segal and Sober’s own account, (P5) is too weak as stated. They require further that F be a ‘substantial’ property, meaning a property that pulls its weight in good scientific generalisations. This rules out properties like being-red-or-weighing-a-hundred-pounds. They admit that they don’t have a particularly clear or informative account of what makes for a substantial property.

(P5) caters specifically for cases in which the supervenience base of F is complex and includes distinguishable micro-properties. This is to allow for cases in which some but not all of the components of the base are involved in the causal transaction at issue. For example: smoking causes cancer. In fact, the activity of smoking involves all sorts of things that do not cause cancer. Only some of what is involved in smoking - the inhalation of material that has certain specific carcinogenic properties - does the work of causing cancer.

Segal and Sober go on to argue that mental causation fits that model: psychological properties mereologically supervene on complex neural (and perhaps

other) properties, some of which are involved in the causal transactions of interest and some of which are not.

Segal and Sober note that their view entails that many second-order functional properties are causally efficacious with respect to the effects relative to which they are defined. This leads them into a dispute with Ned Block. Block allows that second-order functional properties might be efficacious in respect of certain effects in which intelligent beings recognize them. But he denies that they are efficacious in respect of the specific effects that define them. He discusses a bullfighter's cape. Since the cape is red, it is provocative to bulls (let us suppose, he says, for the sake argument, even though it is not in fact true). The bullfighter uses the cape to provoke the bull. Block claims that it is the redness of the cape that caused the bull to be provoked, and not its provocativeness.

Segal and Sober voice the suspicion that Block has confused properties with ways of referring to properties. They draw on the thought that we should not think that a property is inefficacious with respect to a certain effect merely because we pick it out by adverting to that very effect. They discuss the example of solubility, saying this:

we might construe solubility as a second order property, defined thus:

(i) x is soluble iff x possesses some property F such that x 's being F causes x to dissolve when immersed.

But solubility may be defined in other ways. Suppose that there is a unique substantial property that causes objects to dissolve. This property may be defined as follows:

(ii) solubility is the property that causes objects to dissolve when immersed.

So (i) and (ii) provide alternative definitions of the same property. It follows that second order properties may have effects on stupid objects. Of course, there may be no substantial property defined by (ii). But this is what Block would need to show to establish his claim.

But it is Segal and Sober and not Block who display confusion in matters relating to properties and ways of referring to properties. I take it that a unique substantial property that causes objects to dissolve when immersed, would be a micro-property. Definition (ii), if it defines anything real (which I doubt), defines that micro-property. But the existence of such a micro-property is metaphysically contingent. Surely, even if solubility has only one nomologically possible realizer, it has many metaphysically possible ones. But then some metaphysically possible objects are soluble in sense (i) but not in sense (ii), since they have some property that causes them to dissolve when immersed, but not the particular property that causes nomologically possible objects to dissolve. So (i) and (ii) do not define the same property.

And that makes trouble for (P5). For even if a certain range of functionally-specified properties, namely those exemplified by solubility in sense (ii) (if there is such a property) can be efficacious in the relevant respects, (P5) incorrectly licenses the efficacy of others.

Block was right. Functional - that is, genuinely dispositional –properties, whether second-order or not, are not efficacious with respect to their characteristic effects. (For extended discussion of this general issue see Rupert, 2006). To help see this, let us reformulate (i) slightly as (iii):

(iii) *x* is soluble iff (a) *x* would dissolve if immersed and (b) *x* would be caused to do so by some property of *x*'s.

Now, it is part of the notion of a causally efficacious property that it plays a specific role in the causal explanation why its possessors bring about their effects. A property that is efficacious in respect of causing an object to dissolve in water has to enter as a causal factor in the explanation of why the object would dissolve, if immersed in water. Solubility in sense (iii) can't do that because an object's being soluble in that sense just *is*, in part, its being such as to dissolve when immersed.⁸

Segal and Sober repeat Davidson's remark to the effect that statements like 'the cause of *A* caused *A*' are uninformative but, of course, not false. They do this to emphasise the point that just because our description of a cause relates it to an effect does not mean that it is not the cause of that effect. But properties differ from ordinary individual objects and events. Properties can be essentially dispositional, ordinary individuals and events can't. For one thing, the cause of *A* is not even necessarily related to *A*: it might have occurred without *A* having occurred. But an object can't be soluble in sense (iii) and yet not dissolve if immersed. Moreover, the cause of *A* has other properties than merely being the cause of *A*, and its possession of some of those

⁸ I set aside worries about finkish dispositions (Mumford 1996). I am assuming that any fink-resistant analysis of solubility will include an ineliminable counterfactual element of a kind that sustains the claim of this paragraph. Lewis's (1997) account, for example, fits the bill. Lewis himself thinks that dispositions can be efficacious. But his account of this (ibid. 152) employs a very similar magic trick to Kim's. (I suppose Kim got the idea from Lewis). Moreover, he does not engage with the real problem, which is that since an object's possession of a disposition partly consists in its being such that it would cause certain effects in certain circumstances, the object's possession of the disposition cannot be a cause of such effects in such circumstances.

other properties explain why it caused A. And while one can describe solubility in sense (iii) without talking of objects dissolving, one cannot use such descriptions in an explanation of why solubility in sense (iii) causes objects to dissolve. There is no such explanation.

I will say more about dispositions and causal efficacy in the next section. For now, note that (P5) entails that many essentially dispositional properties are causally efficacious in respect of their defining effects; for example solubility, fragility and elasticity, all considered on the model of (iii). So (P5) won't do.

If one wanted follow up on Segal and Sober's start, one could perhaps deal with the problem by adding a further constraint: where F is a substantial and not essentially dispositional property, (P5). Or, indeed, one might try to argue that dispositional properties are not substantial. These paths do not tempt me, however, partly for reasons that will emerge in the next sections and partly because there is another problem with (P5).

Consider the particular sort of red glow that red-hot pokers get.⁹ It is likely that it is a non-strict law that objects with that very specific sort of red glow cause wax to melt. Arguably, the red glow mereologically supervenes on the agitation of the objects' molecules, and it is the latter that causes the agitation of the wax's molecules which, is the supervenience base of the wax's melting. But the red glow is not causally efficacious in respect of melting the wax.

One might doubt that it is a law that objects of that colour melt wax. Maybe a really good special-effects person could create objects that have that colour at room temperature. Maybe. But we can still imagine a possible world in which there are mechanisms that systematically prevent the creation of any such objects, so that there,

⁹ Helen Beebe p.c. suggested (but did not confidently endorse) the counterexample.

it is a *ceteris paribus* law that objects of the requisite colour melt wax. But still it is not because objects have that colour that they melt wax.

Before continuing with the issue of mental causation in particular (section 5, below) I will say more about the metaphysics of properties, dispositions and causation.

5. Properties, Dispositions and Causation

My main concern in this section is to defend a particular general view of the relation between properties and causation, and, in the light of this view, to discuss the role of properties in causal explanation. At issue are all substantial properties that feature in causal explanations. Substantial relations and relational properties, such as siblinghood and reference to Phosphorus are included. The view thus has wide, but not universal scope, failing to apply to properties of abstract objects, such as being a prime number. Causation itself is assumed as a primitive relation.

The view I wish to defend has important features in common with Sydney Shoemaker's and is sometimes attributed to him (erroneously, I think). I will call it 'extreme dispositionalism.' (Hawthorne, 2006 226, calls it 'full-blooded causal structuralism').

The rough idea of extreme dispositionalism is that the properties at issue are just clusters of dispositions. So, for example, what it is for an object to be spherical is for it to be such that if placed on a slope, it would roll down; if placed in certain relations to a mirror it would cast a round image; if placed on plasticine it would create a hemispherical hollow; and so on and so forth.

We can consider a more refined version of extreme dispositionalism which identifies properties not with dispositions but with what Shoemaker (2003 213) calls ‘conditional powers’. Dispositions relate objects to effects in circumstances: to be soluble is to be such that you would dissolve, if immersed, etc.. Conditional powers relativise dispositions (powers) to other properties the object might possess. So, for example, a spherical object has the power to produce a round image, if placed in certain relations to a mirror, if it has a reflective surface, etc.. And it has the power to produce a hemispherical hollow if placed on plasticine, if it is hard and heavy, etc..

As I said, extreme dispositionalism is sometimes attributed to Shoemaker. And, indeed, he does occasionally appear to endorse it: e.g. ‘we can express my view by saying that properties are clusters of conditional powers’ (2003, 213). But it is clear from the context that this is merely an incautious formulation. He explicitly distinguishes properties from dispositions, which he calls ‘powers’. And he thinks of an object’s properties as explaining its powers: an object’s powers are ‘grounded in’ its properties. There is thus a fundamental distinction between Shoemaker’s theory and extreme dispositionalism. According to Shoemaker, properties are not conditional powers, but rather the grounds of those powers: ‘a thing’s powers or dispositions are distinct from, because grounded in, its intrinsic properties’ (2003 213). So, according to Shoemaker, sphericity is not the particular bunch of conditional powers associated with spheres, but rather that property in virtue of which spheres have that particular bunch of conditional powers.¹⁰

So, according to Shoemaker, properties are individuated in terms of the conditional powers that they endow on their possessors. To be spherical, for example, is to have that property in virtue of the possession of which an object has the

¹⁰ This view is consistent with what Hawthorne (2006 226) calls ‘modest structuralism’.

particular conditional powers characteristic of spheres. It is, of course, a consequence of this view that causal relations among properties are metaphysically necessary. On this view, it is metaphysically necessary, for example, that a hard, heavy, spherical etc. object would cause a hemispherical hollow if placed on something soft like plasticine.

Both theories – the extreme one and Shoemaker's – share two excellent features, that I will now discuss.

The first is from Shoemaker (2003 214-5). Suppose, for the sake of reductio, that properties F and G are not individuated by the powers of their possessors. Then you could get two objects that were utterly indistinguishable by any possible test, yet differed in some of their properties. But then, it seems, we could never know anything about these different properties. Of course, it is in principle possible that some good theory might entail the existence of such properties. But there appears to be no reason to believe that there is any such theory. So it is sensible to conclude that $F=G$ iff possessors of F and possessors of G have the same conditional powers, hence that properties are individuated by the powers of their possessors.

This argument is sometimes derided as verificationist. But it is better seen as an application of Occam's razor: do not posit properties that make no difference to anything (cp Hawthorne, 2006 219). Of course, if there were robust properties that failed to associate with causal differences yet served some other explanatory purpose, then there would be reason to believe in them. But there do not seem to be any.

Here is the second desirable feature of the dispositional accounts. Suppose I throw a stone at a window and the window breaks. Why did the window break? Material science tells us about that (see e.g. Gordon, 86). The stone is stiff (that is to say, the material has a high 'Young's modulus', or resistance to deformation). The

stone is massive and spherical. And it is tougher than the glass (it has, *inter alia*, higher impact strength). The stone is travelling fast and a small area of its surface comes into contact with the glass. The glass is thin, relatively weak etc., and it breaks.

Now it is obvious on inspection, I claim, that a number of the predicates featured in the explanation are dispositional and pick out conditional powers. Young's modulus and impact strength are paradigm cases: for the stone to be thus stiff and thus strong *just is*, in part, for it to be such that, given its other properties, it would break such a thing as the window, under suitable impact. So, suppose you have objects somewhat similar to the stone and the window and you throw the former at the latter and the stone breaks, rather than the window. Investigation reveals that their shapes, Young's moduli and every property other than their impact strengths are just the same as the originals. It follows logically that the impact strength differs in at least one case. Either the second window had higher impact strength than the first window, or the second stone had lower impact strength than the first stone, or both.

Now it is tempting to suppose that not all of the properties featuring in the explanation stand in metaphysically necessary connections to the causal powers of their possessors: the stone's shape, for example. It does not seem that being spherical is necessarily connected to the capacity to break windows.

But further reflection shows that the tempting supposition is mistaken. To give a proper account of what having a certain degree, *D*, of strength is, we have to talk about the shapes of strong things. But if we are to treat *D* as a conditional power, thus properly accounted for, then shape is metaphysically implicated in the causal nexus. Strength can only be necessarily connected to the effects of its possessors if shape is too. There is no metaphysically possible world *W* in which the stone has all its actual properties but fails to break the window because sphericity is associated with different

conditional powers in W. It is, therefore, metaphysically necessary that spherical things that share the other properties of the stone, break things such as the window. If there is a world W1, in which the stone doesn't break the window, but is still spherical, then it follows that in W1, either the stone or the window differs from how it is in W in respect of at least one property: either the stone is less stiff or strong or the window is stiffer or stronger, or whatever. It is by the causal relations holding among their possessors that such properties, the dispositional ones like stiffness and strength, are defined.

The connection between, say, fragility and breakages is relatively obvious to us. It is conceptually necessary that fragile things break, when hit hard. Or, to put it better, in a way that makes the conceptual necessity more obvious: it is conceptually necessary that sufficiently fragile things break when hit hard enough by things of the right sort. Material science spells out and quantifies tautologies of that ilk, replacing the crude common-sense notions of fragility and the like with technical, quantitative ones that fit the contours of the real world.

The connection between sphericity and breakages is not obvious to us at first glance. There is nothing in the mere concept of sphericity, considered by itself, that immediately relates it to breakages. However, I do think that the material scientific explanation of what happens when a spherical stone breaks a window articulates a conceptual necessity. Once you understand the concepts involved, you will see that a thing with the properties of the stone would break a thing such as the window, under the relevant circumstances. The explanation includes no appeal to conceptually or metaphysically contingent laws of nature. It just follows from the correct account of the properties of and relations between the two things, that the stone would break the window.

The dispositional theories of properties thus correctly predict that paradigmatic causal transactions, like the breakings of windows by stones, involve metaphysical necessity. This is their second desirable feature.

What seems a virtue to me seems a vice to others. Lewis has objected to the view that causal relations are metaphysically necessary as follows. Lewis adopts a ‘principle of recombination’: (Lewis, 1986 88) ‘roughly speaking the principle is that anything can co-exist with anything else’. The principle expresses a version of the Humean idea that there are no necessary connections between distinct existences. And it naturally leads to the conclusion that laws of nature are not metaphysically necessary, for events that are connected by natural law are distinct existences, and hence are mixed and matched in all different ways across possible worlds. Events of stiff, strong stones hitting windows are distinct from events of windows breaking. Hence there is a possible world where events of the former kind are not followed by events of the latter kind. So the principle of recombination entails that properties featuring in natural laws are not metaphysically connected to the powers of their possessors after all (Lewis 1986 163). Lewis anticipates the possibility that his opponent will simply reject the principle of recombination and responds that he would thereby be leaving the frying pan for the fire. He omits to say which fire.

But there is no clash between the principle of recombination and dispositional theories of properties. There are surely lots of worlds in which people like me throw things like the stone at things like the window and in which the window does not break. We can allow that there are as many of these worlds as Lewis wants. That is not the issue. The issue is how we should describe such worlds. Lewis wishes to describe them as worlds in which stiff, strong stones fail to break windows. The dispositional theorist can just as well describe them as worlds in which the stones or

the windows or both have properties different from their actual ones: the glass, for example, might be stronger or the stone less strong. Hence Lewis's objection misses the mark.

So both dispositional theories look promising. Which is superior? There are three main objections to extreme dispositionalism, which I will now discuss.

The first objection gives voice to a feeling of unease about extreme dispositionalism that I think is widely shared. It is from David Armstrong (Armstrong 1999, p. 67) (who credits a version of it to Richard Swinburne):

Every causal transaction, according to [extreme dispositionalism],¹¹ is a matter of things with certain potentialities bringing it about that these or other things have further potentialities, because properties are analyzed as nothing but potentialities. In Scholastic language, we never get beyond potency to act. Act, so far as it occurs, is just a shifting around of potencies ... 'Where's the bloody horse?' as the poet Roy Campbell might have asked.

The idea seems to be that extreme dispositionalism is viciously regressive or that it leaves properties ungrounded, in some sense. But, as stated, Armstrong's claim amounts to little more than an expression of incredulity. Suppose that I make a snowball. Since I am strong, I can exert pressure on the snow and cause the snowball to be hard. It is very natural to think that my strength and the snowball's hardness are both dispositional properties. So this is a case of one potentiality being involved in the creation of another potentiality. And there is nothing particularly mystifying about that. We understand what is going on. It is a consequence of extreme dispositionalism

¹¹ The actual quote is 'according to Shoemaker'.

that causation is always like that. And why shouldn't it be so? If it can happen some of the time, why should it not happen all of the time? What is needed here is an argument, not a statement of mystification. Of course, extreme dispositionalism is somewhat mystifying. But so is the alternative: that e.g. a fragile object has, in addition to the genuinely dispositional property of (roughly speaking) being such that it would break when hit and the micro-structural property that explains why it would break when hit, a further, intermediate property in which the disposition is 'grounded'.

Extreme dispositionalism might indeed generate an infinite regress: each property is analysed in terms of further properties. But the regress need not be vicious. It would just require there to be infinitely many properties. Alternatively, the theory might generate a circle: the analysis of a property sooner or later leading back to that property itself. But again, the circle need not be vicious (see Holton, 1999 for discussion). In either case, the proposed structure of the metaphysical space of properties is at least coherent. (See Yablo, 1993, for related discussion.)

The second objection can be found in Hawthorne (2006) (a similar argument appears briefly in Armstrong, 1999). Here is a possible causal structure. There are four properties A, B, C and D. A causes C. B causes C. (A&B) cause D. But, the objection goes, there is no way for the extreme dispositionalist to distinguish A and B, since their causal profiles are identical.

There are two ways of spelling the example out, neither of which successfully refutes extreme dispositionalism. The crucial question is whether D is supposed to be caused by the combination of an A object and a B object or whether a single object's possession of A and B is supposed to cause D. If the first, then the extreme

dispositionalist has no problem distinguishing A and B. If the second, then the extreme dispositionalist can reject the possibility.¹²

Let's consider the first option first. On this understanding it is easy for the extreme dispositionalist to distinguish A and B. Suppose that object o1 has A. Then a B object is disposed, in conjunction with o1, to cause D. But an A object is not disposed, in conjunction with o1, to cause D. So A and B correspond to distinct and distinguishable dispositions.

Let us turn to the second option and consider what a world with the alleged structure would look like. Let's consider how objects would behave. Some objects – call these 'F' objects - would cause C. Some objects – call these 'G' objects - would cause D. And that is it. The objector asks us to suppose that in the universe under discussion there is a further, invisible, reality: in fact, the F objects fall into two lots: those that have A and those that have B. Moreover, the G objects that cause D do so because they have both A and B. The difference between A and B, of course, manifests itself in no way at all.

If we allow this possibility, then we open the floodgates. We must also admit that there is a possible universe in which objects behave just like that, but where the F objects divide into three lots: those that have X, those that have Y and those that have Z. The G objects have X and Y, but not Z. There is a third possible universe which is just like that, except that the G objects have X, Y and Z. Then there is a fourth universe where the F objects divide into four lots: those that have W, X, Y and Z. And

¹² I am pretty sure that Hawthorne intends the first option and not the second.

so on, ad infinitum. The extreme dispositionalist can deny that there are all these particular possible universes with a clear conscience.¹³

It is not at all obvious that there is any candidate example of a symmetrical causal structure that won't go the same way: either the mirror properties are dispositionally distinct, or the claim that structure is possible is not independently plausible.

More importantly, though, even if such a structure could be described, that would be of limited interest to debates about the familiar properties in our world whose causal efficacy interests us. Being tough or being spherical or being a belief that Barcelona is beautiful or having a charge of $-(1/3)$ might be genuinely dispositional even if some other quite distinct possible properties are not.

The third objection to extreme dispositionalism comes from Shoemaker. He holds that the manner in which objects can acquire a property can be relevant to the property's individuation. Thus properties cannot be understood purely in terms of dispositions or their grounds. The motivation is due to Richard Boyd. It goes as follows (Shoemaker 2003 232-4):

Imagine a world in which the basic elements include four substances, A, B, C and D. Suppose that X is a compound of A and B and Y is a compound of C and D. We can suppose ... that the property of being made of X and the property of being made of Y share all of their causal potentialities.

¹³ I take it that moderate dispositionalists, like Shoemaker, who hold that properties ground dispositions, would also deny that these are genuine possibilities.

It would follow from extreme dispositionalism that being made of X and being made of Y would be the same property. But Shoemaker finds that counterintuitive, since X and Y would be different substances. So, according to Shoemaker, properties are individuated not just by their associated powers, but also by the circumstances that cause them to be instantiated. His final theory is that property F=property G iff F and G endow the same conditional powers on their possessors and whatever circumstances suffice to cause an instantiation of F suffice to cause an instantiation of G and vice versa.

But there is no compelling reason to suppose that X and Y are distinct substances. It is at least as natural to suppose that there is just the one substance that can be created in either of two ways. Suppose that we lived in Boyd's world. Would we even bother to have two words for X and Y? That is not likely, since it would be a major task to keep track of which samples had which origins. Moreover, we are trying to provide an account of properties that feature in causal explanation. From that perspective, there is no reason to distinguish being made of X from being made of Y.

One can of course distinguish X from Y if one wants to. But the distinguishing properties are the historical ones: having been created from A and B versus having been created from C and D. Those properties are perfectly distinguishable on the extreme causal theory, as long as A, B, C and D are themselves distinct in causal powers. And it is safe to assume that they are, otherwise Boyd's example would make very little sense. There is thus no need to think of X and Y as different substances. They are batches of the same substance that differ only in the manner of their creation.

So extreme dispositionalism survives all three objections. Given its simplicity and elegance, the view stands in good stead and I recommend it.

It is true that both commonsense and science are amenable to the claim that the stone broke the window, in part, because of certain of its properties such as stiffness and impact strength. But, I suggest, the ‘because’ here means *by reason of*. It is the same ‘because’ that features in, for example logic and mathematics: ‘Socrates is mortal because he is a man and all men are mortal.’; ‘Consider a right-angle triangle, with sides A and B of 5 cms and 4 cms respectively. The third side, C, is 3 cms, because $A^2 - B^2 = C^2$ and $5^2 - 4^2 = 3^2$.’¹⁴

The explanation of the why the stone broke the window is rather like the geometrical explanation of why C is 3 cms. And there is nothing wrong with that. It makes transparent to us how various combinations of properties interrelate in the jigsaw puzzles of causal interactions. It has other virtues as well.

One reason why we might want to know, for example, that stiffness and strength are properties of the stone that are relevant to whether it would break a window is that we can identify its possession of those properties without having to see whether it breaks anything. You can feel them with your hand.

Another reason why we might be interested is that while it is conceptually necessary that something as stiff and strong etc. as the stone would break a weaker thing like the window, it is not conceptually necessary that only things with those properties can break windows. Under the right conditions, you could break a window with a large meringue or a mound of jelly. There are many different combinations of properties that can combine to make something able to break a window. So the

¹⁴ Notice that we can also say: consider a right-angle triangle, with sides A and B of 5 cms and 4 cms respectively. Then $A^2 - B^2 = C^2$ because the third side, C, is 3 cms and $5^2 - 4^2 = 3^2$. What was the explanandum has become part of the explanans. We can say that the stone broke the window because it is stiff and strong. And we can just as well say stone is stiff and strong because it would break such a thing as the window.

macroscopic explanation of the event serves to rule out many other possible macroscopic explanations.

Now, if extreme dispositionalism is right, then macroscopic properties in general and psychological properties in particular are dispositional, and so cannot be efficacious in bringing about their characteristic effects. Hence we arrive at epiphenomenalism. But even if extreme dispositionalism is not right, there is a specific reason for supposing that psychological and other macroscopic properties are dispositional. This is the topic of the next section.

5. The Standard Package and the Amazing Coincidence

Let us have another look at The Standard Package and its account of mental causation. It goes as follows. Suppose that it's a law that F events cause G events, where F and G are psychological properties. F has a variety of physical supervenience bases, $m(F)_1, \dots, m(F)_n$. Every time the law is instantiated, the occurrence of the F event is realised by the occurrence of some $m(F)$ event. The $m(F)$ event causes the occurrence of some $m(G)$ event, where $m(G)$ is a supervenience base for G. But if that is right, then it shrieks for explanation. Consider the $m(F)$ s. They all share two properties: (a) they are bases for F and, further, (b) they all tend to cause the occurrence of a base for G. That is a remarkable correlation and there must be some explanation for it. There must be a connection between being a base for F and tending to cause bases for G.

Suppose that you have an account of representation according to which being a mental event with a certain content has nothing to do with what the event is likely to cause. Then you might have a lot of trouble explaining the correlation. That is exactly

the position that Jerry Fodor is in. On his account, the content of a mental representation type is determined by the kinds of thing that would cause tokens of the type to occur. Thus the content of a representation depends on its temporally backward-looking causal sensitivities. Those are evidently distinct from the representation's temporally forward-looking causal powers, what it, in turn, goes on to cause.

Fodor recognizes a version of the problem in Fodor (1994). His account of psychological content entails that it is referential. A person's 'Hesperus' representation and their 'Phosphorus' representations are causally sensitive to the same planet, hence, according to Fodor, they have the same content. But, of course, the representations may behave differently in a person's psychological economy: they have relevantly different forward-looking causal powers. And so Fodor's account leaves open the possibility that psychology will have lots of predictive failures and be no good: for example, a Fodorian reference-only psychology would predict that someone who wants to go to Hesperus and believes that boarding the *USS Morning Star* will get him to Phosphorus, would, ipso facto, be likely to board the *USS Morning Star*.

Fodor spends much time trying to find mechanisms that keep referential psychology viable. These would be mechanisms that prevent Frege cases from happening too often. But he fails (see Segal, 1997 and Aydede and Robbins 2001).

So what could explain the correlation between a mental event's content and its causal powers? There are two kinds of explanation we might consider: metaphysically necessary and metaphysically contingent. Fodor looked for a metaphysically contingent one. And one would expect any account of content that relies entirely on

historical facts to be in the same position and stand in need of supplementation with an account of contingent mechanisms that explain the correlation.

But obviously functionalism has a built-in explanation of the correlation, one that depends entirely on the metaphysics of content. What it is to be an event with a particular content is (in part) to have certain specific causal powers. So, for example, it is no accident that my desire to eat an Italian sausage tends to cause me to eat an Italian sausage, under certain conditions. Having that tendency is a metaphysically necessary condition for being a desire with that content. So, of course, any micro-properties that realise the desire will tend to cause events that realise my eating an Italian sausage.

Interpretationism is afloat in the same boat. According to interpretationism, a set of states of a physical system is representational iff there exists an interpretation of those states such that actual (and counterfactual) complex behaviour of the system consistently makes (and would make) reasonable sense under that interpretation. The notion of interpretation involved is the logical one: there is a mapping from the states onto some suitable domain. According to interpretationism, the states then really represent what they represent in the logical sense, under any sensible interpretation.

Interpretationism can explain the correlation between a state's content and its effects. Being sensibly interpretable is a holistic property of a system. An interpretation of a given state will only make sense in the context of an appropriate interpretation of the state's causes and effects. And the relevant kind of appropriateness is just what is required to explain the correlation. For example, a given state will only be interpretable as, say, a desire to eat an Italian sausage if, under certain circumstances, it causes a state interpretable as an eating of an Italian sausage.

Interpretationism is not usually thought of as a variety of functionalism. But (as leading interpretationist, Robert Cummins, admits (Cummins 1989)) they are metaphysically equivalent. Whether a particular interpretation makes sense of a system depends only on causal relations among states of the system.¹⁵ So I will use ‘functionalism’ to include interpretationism in its extension.

Functionalism’s capacity to explain a correlation that cries out for explanation, while other accounts of content leave the mystery unexplained, seems to be a strong consideration in favour of it.¹⁶ I will proceed on the assumption that some variety of functionalism is true.¹⁷

¹⁵ Of course, there are many different varieties of functionalism. Some functionalists take causal relations between the system and the external world to be relevant to content (e.g. Harman 1987). But then some interpretationists have a parallel constraint on the set of true interpretations (e.g. Gallistel 1993).

¹⁶ I don’t believe that functionalism is true for phenomenal properties. If phenomenal properties feature in causal laws, then perhaps they reduce to neurological properties. They might well be much less multiply realisable than representational properties and species-specific reductions might be viable. Henceforth I will consider only representational properties.

¹⁷ Teleological theories of content also have resources that might help to explain the coincidence. According to teleological theories of content, the content of a representation is explained in terms of the representation’s function. And the function of a representation will typically have a lot to do with what it causes in what circumstances. As you almost certainly know, frogs have a mechanism that is triggered by small objects whizzing past that causes an automatic tongue protrusion. If the flying object is a fly within range, then the frog catches it with its tongue and eats it. Let us suppose that the mechanism is triggered by a representation with the content: *fly*. The reason the representation means *fly* is because it causes the frog to do something that is beneficial for it in the presence of flies. If we put in *fly* for F and snapping for G in the *prima facie* coincidental structure, then we see that is not a coincidence that m(F)s represent flies and cause tongue protrusions. But teleological theories have grave problems, see in particular Peacocke (1992 129-33).

6. Conclusion: Epiphenomenalism Versus Overdetermination

If functionalism is true, then mental properties are dispositional, hence not causally efficacious, and epiphenomenalism is true. But mental properties are none the worse for that. They are just as real as paradigmatic physical properties like stiffness and impact strength. And they can play indispensable roles in excellent causal explanations, just like their physical counterparts. There is nothing wrong with or second-rate about dispositional properties.

I said above that I would present my view as a variety of epiphenomenalism. It seems an appropriate label. But notice that it is very similar to another view. Suppose we replace the notion of a causally efficacious property with, say, the notion of a causally explanatory property.¹⁸ We could explicate this notion by examples; the stone's stiffness and impact strength are causally explanatory in respect of its breaking the window, its colour and age are not. To be causally explanatory is to feature in a causal explanation in the familiar way. It is just that we now understand that that familiar sort of explanation has a different character than we thought. We thought the properties were only nomologically related to the effects they explain, whereas actually they relate by metaphysical necessity. We thought the 'because' in the 'the stone broke the window because it was strong ...' was an *affect* because, whereas actually it is a *reason* because.

¹⁸ Compare Jackson and Pettit (1988) on what they call 'causally relevant' properties. These are relevant to causal explanation even if causally inefficacious.

We might then use ‘epiphenomenalism’ to apply to the thesis that mental properties are not causally explanatory in respect of the effects of their possessors. Such terminological moves seem not unreasonable. We could then reasonably claim to be proponents of an overdetermination approach rather than of epiphenomenalism. It appears that the difference between overdetermination and epiphenomenalism is largely rhetorical.

Philosophy
King’s College, London

BIBLIOGRAPHY

- Armstrong, D. 1999: The causal theory of properties: properties according to Shoemaker, Ellis and others, *Philosophical Topics* 26, 25–37.
- Aydede, M., and Robbins, P., 2001: Are Frege cases exceptions to intentional generalization? *Canadian Journal of Philosophy* 31 1, 1-22.
- Bennett, K., 2003: Why the exclusion problem seems intractable and how, just maybe, to tract it. *Nous*, 37 3, 471-497.
- Block, N., 1990: Can the mind change the world. In G. Boolos, (ed.), *Essays in Honor of Hilary Putnam*. Cambridge: Cambridge University Press.
- Block, N., 2003: Do causal powers drain away? *Philosophy and Phenomenological Research*, 67: 133-150.
- Anscombe, E., and Geach, P., (eds.) 1971: *Descartes: Philosophical Writings* Norwich: Fletcher and Son.
- Cummins, R., 1989: *Meaning and Mental Representation*, Cambridge, MA.: MIT Press.
- Davidson, D. 1970: Mental events. In L. Foster and Swanson, J. (eds.), *Experience and Theory*. Amherst: University of Massachusetts Press.
- Fodor, J., 1974: Special sciences. *Synthese*, 28: 97-115.
- Fodor, J., 1994: *The Elm and The Expert*. Cambridge, MA.: MIT Press.
- Gallistel, C., 1993: *The Organization of Learning* Cambridge, MA.: MIT Press
- Gordon, J., 1968: *The New Science of Strong Materials or Why You Don't Fall Through the Floor*, New Jersey: Princeton University Press.
- Haggard, P., and Eimer, M., 1999: On the relation between brain potentials and the awareness of voluntary movements. *Experimental Brain Research*, 126: 128–133.

- Harman, G. 1987: (Non-solipsistic) conceptual role semantics. In E. Lepore (ed.), *New Directions in Semantics*, London: Academic Press.
- Hawthorne, J., 2006: Causal structuralism. In his *Metaphysical Essays*, Oxford: Oxford University Press. Originally published in *Philosophical Perspectives* 2001 15: 361-378.
- Holton, R., 1999: Dispositions all the way round. *Analysis*, 59: 9-14.
- Jackson, F., 1982: Epiphenomenal qualia. *Philosophical Quarterly*, 32: 127-36.
- Jackson, F. and Pettit, P., 1988: Functionalism and broad content. *Mind* 97: 381-400.
- Kim, J., 1984: Epiphenomenal and supervenient causation. In P. French *et. al.* (eds.) *Midwest Studies in Philosophy*, IX. Minneapolis: University of Minnesota Press.
- Kim, J., 2000: *Mind in a Physical World*, 1st paperback edition, Cambridge, MA.: MIT Press.
- Kim, J., 2005: *Physicalism or Something Near Enough*, Princeton and Oxford: Princeton University Press.
- Lewis, D., 1997. Finkish dispositions. *The Philosophical Quarterly* 47: 143-158.
- Lewis, D., 1980: Mad pain and Martian pain. In N. Block (ed), *Readings in the Philosophy of Psychology*, vol. 1, 216-22. Cambridge MA.: Harvard University Press.
- Lewis, D., 1986. *On the Plurality of Worlds* Oxford UK and Cambridge MA. Blackwell.
- Mumford, S., 1996: Conditionals, functional essences and Martin on dispositions. *The Philosophical Quarterly* 46: 86-92.
- Papineau, D., 2000: The rise of physicalism. In J. Wolff and Stone, M., (eds) *The Proper Ambition of Science*, London: Routledge.
- Peacocke, C., 1992. *A Study of Concepts* Cambridge MA.: MIT Press.

- Rupert, R., 2006: Functionalism, mental Causation, and the problem of metaphysically necessary effects. *Noûs* 40: 256-83
- Segal, G., 1997: Content and computation: a critical notice of Jerry Fodor's *The Elm and the Expert*. *Mind and Language*, 6, 491-501.
- Segal, G., and Sober, E., 1991: The causal efficacy of content. *Philosophical Studies*, 63, 1-30.
- Sider, T., 2003: What's so bad about overdetermination? *Philosophy and Phenomenological Research* 67: 719-726
- Yablo, S., 1992: Mental causation. *Philosophical Review*, 101: 245-280.
- Yablo, S., 1993: Definitions, consistent and inconsistent. *Philosophical Studies* 72, 147-75.